

The ISO Problem: Decentralized Stochastic Control via Bidding Schemes

Rahul Singh, P. R. Kumar and Le Xie

Abstract—We consider a smart-grid connecting several agents, modeled as stochastic dynamical systems, who may be electricity consumers/producers. At each discrete time instant, which may represent a 15 minute interval, each agent may consume/generate some quantity of electrical energy. The Independent System Operator (ISO) is given the task of assigning consumptions/generations to the agents so as to maximize the sum of the utilities accrued to the agents, subject to the constraint that energy generation equals consumption at each time.

This task of coordinating generation and demand has to be accomplished by the ISO without the agents revealing their system states, dynamics, or utility/cost functions. We show how and when a simple iterative procedure converges to the optimal solution. The ISO iteratively obtains electricity bids by the agents, and declares the tentative market clearing prices. In response to these prices, the agents submit new bids.

On the demand side, the solution yields an optimal demand response for dynamic and stochastic loads. On the generation side, it provides the optimal utilization of stochastically varying renewables such as solar/wind, and generation with fossil fuel based generation with dynamic constraints such as ramping rates. Thereby we solve a decentralized stochastic control problem, without agents sharing any information about their system models, states or utility functions.

I. INTRODUCTION

We consider the problem faced by the electricity grid operator, called the Independent System Operator (ISO). If the ISO knows the total demand of the loads, it has to solve the problem of allocating the required power among different generators so that the total cost of production is minimized¹. This problem can be solved by the generators bidding their marginal cost curves, and the ISO performing the optimization to obtain and declare the market clearing price. There are also additional aspects such as assuring that the power flow can be delivered over the network [1]–[3].

The above deterministic static model with a fixed demand is insufficient for the oncoming era when we want to maximize the integration of renewable energy sources such as wind and photo voltaic, which are dynamic and vary unpredictably with time. Their modeling requires a

dynamic stochastic system. Dynamic models can also be used to model features such as ramping constraints that are important for modeling fossil fuel generators.

When employing renewable energy we need to make the demand adapt to the availability of renewable energy, called “demand response”, in contrast to the traditional scenario where demand is inflexible and supply needs to match whatever demand is. Loads generally have dynamic constraints, since, for example, air conditioners can be deferred for a while but not indefinitely. So loads also need to be modeled as dynamic control systems. Further, since environmental variables such as temperature are involved, they may be uncertain, and hence will also generally need to be modeled as stochastic dynamic systems.

Such dynamic models can also model storage devices where the state is the amount of energy stored, and prosumers, such as homes with solar panels, which may switch at uncertain times from being consumers to generators. Therefore we model all the agents involved as stochastic dynamical systems.

The goal in operating this system is to maximize the sum of the utilities of all the agents. There are however several constraints on information sharing that need to be respected. Individual agents may be averse to sharing system states with each other. They may not even be willing to share their individual system models or their individual utility functions for several reasons ranging from the competitive nature of commercial enterprises, to protecting privacy of consumers’ home states.

The overall systemwide optimality of such a system is sought to be achieved by the ISO. It needs to both determine the optimal demand response over time, as well as allocate generation over time among the lowest cost generators, in the face of stochastic uncertainty, and to do so at minimum systemwide cost. The ISO would like to achieve this by simply determining prices and leaving each agent to its own selfish utility maximization, as in general equilibrium theory [4].

Our main results are the following. We establish iterative interaction processes such as tatonnement [5] under which the ISO can indeed perform this task for certain stochastic dynamic systems. We also address the complexity of this task under several scenarios. In the case where the agents can be modeled as linear Gaussian stochastic systems and the cost functions are quadratic, we show that a simple scheme not involving contingent markets yields the systemwide global optimum.

This material is based upon work partially supported by NSF under Contract Nos. ECCS-1546682, CPS-1239116, and Science & Technology Center Grant CCF-0939370.

Rahul Singh, P. R. Kumar and Le Xie are with Texas A&M University, College Station, TX 77840, USA. rsinghl@tamu.edu, prk@tamu.edu, lxie@tamu.edu

¹There are additional aspects such as security against contingencies, etc., that we neglect here.

II. SYSTEM MODEL

Consider a smart-grid consisting of M agents which may be generators, loads, prosumers or storage devices, each modeled as a stochastic dynamical system. The following are the key ingredients of our system:

- 1) *Randomness* is modeled through a probability space $(\Omega, \mathcal{F}, \mathbb{P})$. The “state of the world” $\omega \in \Omega$, and captures “random” phenomena such as weather, wind-speed, coal shortage, or a damaged wind-turbine. It affects agent i through the random processes $N_i(\omega, t)$ and $N_c(\omega, t)$ $t = 0, 1, 2, \dots, T-1$. We regard $N_c(t)$ as a “common” uncertainty that affects and is known to causally by all agents (e.g., temperature of a city), while $N_i(t)$ is a “private” uncertainty specific and known causally only to agent i . We denote by \mathcal{F}_t the sigma algebra generated by all the noises up to and including time t .
- 2) *Agents*: Each agent i has a state $X_i(t)$ at time t known to it, that evolves as,

$$X_i(t+1) = f_i^t(X_i(t), U_i(t), N_i(t), N_c(t)), \quad (1)$$

where $U_i(t) \in \mathbb{R}$ is the amount of electricity supplied (negative if consumed) to the grid by agent i at time t . Each discrete time instant corresponds to a 15-minute interval of the real-time market implemented by the ISO.

- 3) *One-step Cost function* of an agent i , $c_i(x_i, u_i)$ (or its negative, a one-step utility function $-c_i(x_i, u_i)$). For producers, this cost could be due to labor, coal, etc.. For consumers, this could represent the cost incurred due to the high temperature of house/business facility, or the cost incurred due to a delay in performing a task resulting from non-purchase of electricity.
- 4) *System Operating Cost* is the expected value of the sum of the finite horizon total costs incurred over the time duration $\{0, 1, 2, \dots, T\}$ by all the agents,

$$\mathbb{E} \left(\sum_{t=0}^T \sum_{i=1}^M c_i(X_i(t), U_i(t)) \right). \quad (2)$$

The time horizon T can, for example, be 96 which corresponds to one day. It is the total electricity generation cost minus the utility provided to the consumers.

- 5) *Energy Balance Constraint*: The basic constraint that we will focus on is that total generation must equal total consumption at each time t , i.e., $\sum_{i=1}^M U_i(t) = 0$ at each time t .
- 6) *The ISO* is an agency that accepts electricity purchase/sale bids that are submitted by the agents for each time slot $t = 0, 1, \dots, T$. In our model, we allow for the agents and ISO to iterate on the bids before the market clearing price is declared. Once the iterations have converged, the ISO declares the market clearing prices, and the agents

generate/consume the agreed electrical energies at the declared prices.

III. THE ISO PROBLEM

The ISO problem is to solve the following constrained stochastic dynamic control problem,

$$\min \mathbb{E} \left\{ \sum_{t=0}^T \sum_{i=1}^M c_i(X_i(t), U_i(t)) \right\}$$

$$\text{such that } \sum_i U_i(t) = 0, t = 0, 1, \dots, T-1,$$

$$\text{and } X_i(t+1) = f_i^t(X_i(t), U_i(t), N_i(t), N_c(t)), \text{ for } i = 1, 2, \dots, M, \text{ and } t = 0, 1, \dots, T-1.$$

(ISO Problem)

The expectation above is taken with respect to the combined uncertainty or “noise” process $N(t) := (N_1(t), N_2(t), \dots, N_M(t), N_c(t))$ for $t = 0, 1, \dots, T-1$.

The noises $\{N_1(t), N_2(t), \dots, N_M(t), N_c(t)\}$ may be dependent random variables. There can also be dependence across time.

IV. PROBLEM STATEMENT, KEY QUESTIONS AND GOALS

The key issues of interest are the following:

- 1) Since the ISO Problem is a decentralized control problem with non-classical information structure [6], [7], we would like to know whether it is possible to achieve the exact optimal performance as attained by centralized control. As our analysis will show, the ISO is able to attain optimal coordination amongst the M dynamic systems through announced “prices” under some conditions.
- 2) *Decentralized Optimization and Sufficient Statistics* The ISO Problem can be viewed as a constrained Markov Decision Process (MDP) [8]. The current state of our knowledge does not allow us to handle general MDPs with different observation patterns and different cost functions. Our scheme establishes certain “sufficient statistics” for each agent to make optimal decisions in a decentralized environment [9], [10].

We will show that there exist simple “iterative bidding schemes” (IBS) which yield the same performance as that of the optimal centralized controller under some models.

V. RELATED WORKS

No similar results appear to be known for the general decentralized stochastic control problem. Team problems have been extensively studied, e.g., [7], [11], but the formulations are very restrictive in that each agent needs to know the system dynamics of the other agents. Even when the models are known, there are still considerable difficulties in decentralized stochastic control. When agents do not share observations, severe complexity can set in, even in an otherwise linear quadratic Gaussian problem, as pointed out by Witsenhausen in

his counterexample of a two stage problem [6]. The role of observation, signaling [7], and the trade-off between communication and control, are evident from this counterexample [6]. There are some recent structural results [12] and on sufficient statistics [13] under the restrictive assumption that the agents know the dynamics of all other agents. But the proposed solutions suffer from the curse of dimensionality.

We show below that the ISO Problem is an example of a decentralized control system with non-classical observation patterns in which signaling can successfully result in globally optimum performance. The agents need not reveal their observations, state values, their individual system dynamics or their individual cost-functions. We construct concrete signaling schemes which encode-decode the information required in order to recover the same performance as that of centralized control. From the economics side, this work is an extension of general equilibrium theory [14]. To the authors' knowledge there does not appear to be any similar result for coordinating multiple LQG systems or the efficiency of the simplified signaling.

In the setting of the energy market, [15], [16] discuss a general framework for the operation of an electricity grid via the establishment of a marketplace in which the electricity purchases are done on the basis of spot prices. Thus, electrical energy is treated as a commodity which can be bought, sold, and traded, taking into account its time-and space-varying values and costs, and a framework is presented for the establishment of an energy marketplace. Our results can be viewed as providing guidelines for the operation of this marketplace via simple bidding schemes when the agents are nonlinear stochastic dynamical systems. Our scheme maximize the net social utility.

Viewed from the power system end, there have been many efforts since the deregulation of the electricity sector on a market-based framework to clear the system. Ilic et al. [17] have proposed a two-layered approach that internalizes individual constraints of market participants while allowing the ISO to manage the spatial complexity. The approximated MPC algorithm is shown to perform well in many realistic applications.

In order to analyze the strategic interactions between the ISO and market participants, game theoretical approaches have been proposed, e.g., Zhu et al [18] uses a Stackelberg game framework for studying economic dispatch with demand response. Wang et al [19] pose the problem as a noncooperative game, and prove the existence of a Nash equilibrium under some assumptions.

One of the major challenges we address is how to elicit optimal demand response without revealing the inherent dynamic models of the loads to the ISO.

VI. A TREE VISUALIZATION OF SYSTEM RANDOMNESS

A tree visualization of the system randomness will be insightful in the discussions to follow; see Fig 1. The combined system comprising of all the M agents evolves as,

$$X(t+1) = f^t(X(t), U(t), N(t)). \quad (3)$$

Let us assume for the time being that the noise process $N(t)$ is allowed to assume only finitely many values at each time. We can then construct an uncertainty tree of depth T , in which the root node corresponds to initial system state, and every path from the root to a leaf node corresponds to a unique realization of the noise sequence $(N(0), N(1), \dots, N(T-1))$.

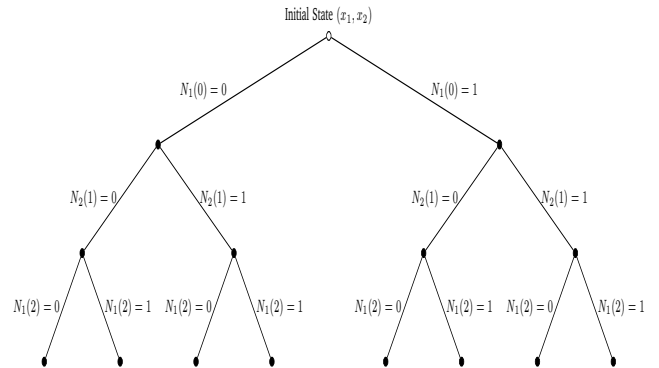


Fig. 1. A Tree based visualization of randomness for a two agent system evolving over three bid times. The noise values are allowed to be binary, 0 or 1.

VII. ITERATIVE BIDDING SCHEMES

The key contribution of this work is to propose solutions to the ISO Problem in the form of Iterative Bidding Schemes (IBS), as in Walrasian tatonnement [4]. Such schemes intertwine two simple processes, which we call *Bid Update* and *Price Update*. We begin by defining two key elements of the IBS, the bid function and the price function.

Bid Function: A bid sequence by agent i specifies to the ISO how much electricity that agent will purchase (negative if supplying) in every time period from that time till the final time. At time t it is a sequence of the form $(U_i(t), U_i(t+1), \dots, U_i(T))$. A bid function (in short just “bid”) specifies the bid sequence, as a function of the past history of observed noise $N(s), s < t$. In Figure 1 the bid function of each agent simply specifies, for each node in the tree, the amount of electricity that agent i is willing to purchase when the system passes through that node if it ever does so. We note that $U_i(\omega, t)$ is adapted to the filtration \mathcal{F}_t . The bid function of agent i will be denoted by U_i .

A **price function** is a function announced by the ISO, which specifies for each time t , as a function of the past history of noise $N(s), s < t$, the price $\lambda(t)$ at

which electricity will be sold/bought in the market. In the tree example of Figure 1, this corresponds to the market clearing price corresponding to each node of the tree. The price function $\{\lambda(\omega, t)\}$ is also an \mathcal{F}_t -adapted stochastic process, which will be denoted by λ .

Bid Update: Let us suppose, for the time being, that the ISO has somehow declared a price function λ . In the *Bid Update*, each agent i changes its bid in response to the price function λ . In order to derive its new bid, it solves the following problem, dubbed Agent i 's Problem,

$$\min \mathbb{E} \left\{ \sum_{t=0}^T c_i(X_i(t), U_i(t)) + \lambda(t)U_i(t) \right\}.$$

It maximizes agent i 's total net utility, defined as the utility $-c_i(X_i(t))$ it derives from its state being $X_i(t)$, minus the amount $\lambda(t)U_i(t)$ it pays for the electricity.

Price Update The ISO updates the price function in response to the agents having submitted their updated bids. Since in our context here the sole purpose of the ISO is to make sure that the net demand equals net supply, we will consider a simple rule by which it raises prices if demand exceeds supply and reduces otherwise, i.e., guided by the excess consumption function. Suppose the previous price was λ^k and the bid was U^k . Then, the Price Update is,

$$\lambda^{k+1}(t) = \lambda^k(t) (1 - \alpha_k) + \alpha_k \left(\sum_i U_i^k(t) \right)$$

where $\alpha_k > 0$ is an ‘‘adaptation gain’’. We employ the choice $\alpha_k = 1/k$, which satisfies the twin conditions $\sum_{k=0}^{\infty} \alpha_k = \infty$, and $\sum_{k=0}^{\infty} \alpha_k^2 < +\infty$, a common convergence condition in stochastic approximation [20].

It will be the object of the following section, to show that an iteration of Bid Update-Price Update can solve the ISO Problem under some conditions.

VIII. THE DETERMINISTIC CASE

First we consider the ISO Problem for deterministic systems,

$$\begin{aligned} & \min \sum_{t=0}^T \sum_{i=1}^M c_i(x_i(t), u_i(t)) \\ & \text{such that } \sum_i u_i(t) = 0, \text{ for } t = 0, 1, \dots, T, \\ & \text{and } x_i(t+1) = f_i^t(x_i(t), u_i(t)), \text{ for} \\ & i = 1, 2, \dots, M, \text{ and } t = 0, 1, \dots, T-1. \end{aligned} \quad (4)$$

The intermediate variables $x_i(t)$ can be expressed in terms of the inputs $u_i := (u_i(1), u_i(2), \dots, u_i(T-1))$ and thus the cost term $\sum_{t=1}^T \sum_{i=1}^M c_i(x_i(t), u_i(t))$ can also be expressed solely as a function of the inputs $u_i, i = 1, 2, \dots, M$. Convexity plays a major role, as noted by Arrow [4].

Assumption 1 (Convexity Assumption): For $i = 1, 2, \dots, M$, the function $\sum_{t=1}^T c_i(x_i(t), u_i(t))$ is convex in the input vector $(u_i(1), u_i(2), \dots, u_i(T-1))$.

We will now derive a solution to the ISO Problem under Assumption 1, and show that it achieves the same performance as that of optimal centralized control.

The cost is convex in the vectors u_i . Employing the definition of each $x_i(t)$ as $f_i^t(x_i(t-1), u_i(t-1))$, the associated Lagrangian and dual function are given by,

$$\begin{aligned} \mathcal{L}(u, \lambda) &:= \sum_{i=1}^M \left\{ \sum_{t=0}^T c_i(x_i(t)) + \lambda(t)u_i(t) \right\}, \\ D(\lambda) &:= \min_u \mathcal{L}(u, \lambda), \end{aligned}$$

where $u := (u_1, u_2, \dots, u_M)$, and $\lambda := (\lambda(0), \lambda(1), \dots, \lambda(T-1))$. The Lagrangian is the sum of the costs incurred by each individual agent. Hence, given the Lagrange multipliers λ , the inputs u_i minimizing the Lagrangian can be calculated in a decentralized fashion, with each agent i solving its own problem called Agent i 's Problem

$$\begin{aligned} & \min \sum_{t=0}^T c_i(x_i(t)) + \lambda(t)u_i(t), \\ & \text{subject to } x_i(t+1) = f_i^t(x_i(t), u_i(t)) \\ & \text{for } t = 0, 1, \dots, T-1. \end{aligned} \quad (5)$$

Each agent i then submits this optimal $u_i(\cdot)$ to the ISO as its bid. This enables the computation of the dual function for each value of λ .

Note that the sub-gradient with respect to λ of the Dual function $D(\lambda)$ is $(\sum_i u_i^k(0), \sum_i u_i^k(1), \dots, \sum_i u_i^k(T-1))$. Since the dual problem of finding the prices λ that maximize $D(\lambda)$ is convex, it can be solved via the sub-gradient iteration [21]–[23].

$$\lambda^{k+1}(t) = \lambda^k(t) (1 - \alpha_k) + \alpha_k \left(\sum_i u_i^k(t) \right), t \geq 0, \quad (6)$$

where k is the index which keeps track of the iteration number. The iterations end when the price vector $\lambda(t)$ converges to the optimal value $\lambda^*(t)$. The resulting solution is optimal for the ISO Problem due to the convexity assumptions.

IX. PRIVATELY OBSERVED NOISE

Suppose that even though the agents do not observe the private noises of other agents, or know their system dynamics or utility functions, they know the laws of the combined noise process, $\mathcal{L}(N(t))$. That is, in the context of the uncertainty tree of Section VI, the agents know the topology of the tree, and the transition probabilities along the edges.

We will now show that under the following convexity assumption, the ISO Problem has an optimal solution.

Assumption 2: The function

$$\sum_t c_i(X_i(t), U_i(t)), i = 1, 2, \dots, M, \quad (7)$$

is convex in the vector $\{U_i(t), t = 0, 1, \dots, T-1\}$ for fixed noise sequence.

The algorithm presented is iterative, and composed of Bid-Price updates. The bid submitted by each agent i at time s is a random process that maps the space $\Omega \times \{s, s+1, \dots, T-1\}$ to \mathbb{R} . This is akin to Arrow's [4] approach of treating each good available at a certain time and place as a separate good. The bid process is adapted to the filtration \mathcal{F}_t . At each time t , it specifies to the ISO, as a function of the past noise $N(s), s < t$, the amount of electricity that the agent is willing to purchase at time t . Only the initial bid, $U_i(s)$ is implemented at time s , as in Model Predictive Control. Figure 2 summarizes the algorithm.

Theorem 1: Algorithm 1 solves the ISO Problem when the cost functions satisfy Assumption 2, with each agent i having access only to its private noise $N_i(t)$ and the common noise $N_c(t)$, with the law of the combined noise process, i.e., $\mathcal{L}(N(t))$, being known publicly.

Proof: Let us first consider the special case where there is only a commonly observed noise N_c and there are no private noises, called the Commonly Observed Noise Problem. Suppose for simplicity of visualization the noise process $N(t)$ assumes only finitely many values allowing it to be represented by a tree as in Fig. 1.

Let us suppose that $x(0)$ is fixed, without loss of generality. Let p_v denote the probability of node v in the uncertainty tree. The depth of the node in the tree indicates time. Every Markov policy, mapping states and time to actions, specifies an action $U(v) := (U_0(v), U_1(v), \dots, U_M(v))$ satisfying $\sum_i U_i(v) = 0$ for every node v in the tree. This is easily seen by recursion starting at the root which corresponds to the initial time and state of the system, and noting that each node then also indicates the state of the system at that time. Now consider also a more general "tree policy" that specifies a $U(v) := (U_0(v), U_1(v), \dots, U_M(v))$ satisfying $\sum_i U_i(v) = 0$ for every node v in the tree. It is more general than a Markov policy since two nodes in the tree at the same depth may correspond to the same state $X(t)$ but a tree policy is allowed to prescribe different actions for them. Hence the class of tree policies also contains an optimal policy.

For every such tree policy, for every node v , there is a unique sequence of actions $U^v := \{U(0), U(1), \dots, U(t)\}$ that was taken in the preceding t steps, where t denotes the depth of the node v . The state $X(t)$ at time t corresponding to the node v is thereby determined by (v, U^v) . The centralized optimization problem can then be written as the following optimization problem,

$$\begin{aligned} \min & \sum_{i=1}^M \sum_v p_v c_i(v, U^v) \\ \text{such that} & \sum_i U_i(v) = 0, \forall v. \end{aligned}$$

Note that $c_i(v, u)$ is convex in u . Hence this is a convex

programming problem with no duality gap. Associating Lagrange multiplier $\lambda(v)$ with the constraint $\sum_i U_i(v) = 0$, and letting $\lambda := \{\lambda(v)\}$, we obtain,

$$\mathcal{L}(U, \lambda) := \sum_{i=1}^M \sum_v p_v \left\{ \sum_v c_i(v, U^v) + \lambda(v) U_i(v) \right\}.$$

We will call the process $\lambda(v)$ as the "price process". Each agent submits a bid for each possible partial realization v of the noise process, while the ISO specifies a price at each v . Now the proof parallels the proof in the deterministic case.

This proof extends to the case where there are also private noises. At time 0, there are no private noises $N_i(-1)$, and so the above proof holds at time 0. Noting this, it follows that the result also holds at each time $s \geq 1$ since the bid-price iteration is repeated at each such time, and we can simply regard s as the new "initial" time. ■

Algorithm 1

Assumption: The law of the combined noise process $\mathcal{L}(N)$ is common knowledge of all agents and ISO.

for bidding times $s = 0$ to $T-1$ **do**
 $k = 0$

repeat

Each agent i solves the problem

$$\min \mathbb{E} \left\{ \sum_{t \geq s} c_i(X_i(t), U_i(t)) + \lambda^k(t) U_i(t) \right\},$$

(Agent i 's Problem)

with initial condition $X_i(s)$ for the optimal $\{U_i^k(t), s \leq t \leq T-1\}$, and submits it to ISO.

ISO declares new price, i.e.

$$\lambda^{k+1}(t) = \lambda^k(t) (1 - \alpha_k) + \alpha_k \left(\sum_i U_i^k(t) \right),$$

for times $t \geq s$.

$k \rightarrow k+1$

until $U_i^k(t)$ converges a.s. to $U_i^*(t)$, $\forall t \geq s$
ISO implements $U_i^*(s)$

end for

We note that the assumption of the common knowledge of $\mathcal{L}(N(t))$ can be removed by utilizing the technique of Stochastic Approximation or other learning techniques [20], [24], [25], so that the agents can "learn" the laws $\mathcal{L}(N(t))$.

The above algorithm above is exponentially complex in T due to the number of possible states in the tree, even when there are only two possible values for each $N_i(t)$. In the next section we will see that we can dramatically simplify the algorithm in the LQG context.

X. THE CASE OF LINEAR SYSTEMS

This section treats the special case of the ISO Problem when the M agents have linear Gaussian dynamics and quadratic costs. The noises of all agents are independent and mean zero. Each agent i has a quadratic cost criterion, i.e., the cost functions $c_i(x_i, u_i) = x_i^\top Q_i x_i + u_i^\top R_i u_i$ are quadratic, with weighting matrices $Q_i \geq 0$ and $R_i > 0$. Let us call this the Distributed Constrained LQG (DCLQG) Problem:

$$\begin{aligned} \min \mathbb{E} & \left(\sum_{t=1}^T \sum_{i=1}^N X_i^\top(t) Q_i X_i(t) + U_i^\top(t) R_i U_i(t) \right) \\ \text{subject to } & X_i(t+1) = A_i X_i(t) + B_i U_i(t) + B_i N_i(t), \\ & t = 0, \dots, T-1, \\ \text{and } & \sum_i U_i(t) = 0, t = 0, \dots, T-1. \end{aligned} \quad (8)$$

(The case of time-varying systems is analogous to time-invariant systems, and omitted for brevity) We will assume that the system dynamics given by (A_i, B_i) , the cost functions given by (Q_i, R_i) , and the observation structure are all private, i.e., none of the agents have knowledge of the system parameters or the costs of the other agents, and that the state process X_i is observed only by the agent i .

We will derive an Iterative Bidding Scheme which is much simpler than the algorithm proposed in Section IX in the following critical aspect: The bid function submitted at time t specifying the quantity of electricity that agent i is willing to purchase at times $t, t+1, \dots, T-1$ does not depend on the outcomes of noise sequence $N(s), s \geq t$. It is simply a vector $(u_i(t), u_i(t+1), \dots, u_i(T-1))$ comprising of $T-t+1$ entries. This is a drastic reduction in complexity of the bidding scheme. At each time t , the following iteration takes place: Each agent bids a vector of future purchases corresponding to a deterministic certainty equivalent system, in response to prices announced by the ISO for future power, and the ISO updates the prices in return, until convergence.

Definition 1 (Certainty Equivalence): A stochastic control problem is said to possess the property of certainty equivalence if the optimal policy for the stochastic control problem coincides with the optimal policy for the corresponding deterministic control problem in which the noise is absent.

Theorem 2: The following bidding scheme achieves optimality for the ISO Problem with LQG agents. At each time s , in response to the k -th iterate of the price sequence $(\lambda^k(s), \lambda^k(s+1), \dots, \lambda^k(T))$, agent i announces the optimal open loop sequence $(u_i^k(s), u_i^k(s+1), \dots, u_i^k(T))$

for the deterministic LQ problem:

$$\begin{aligned} \min & \sum_{t \geq s} x_i^\top(t) Q_i x_i(t) + u_i^\top(t) R_i u_i(t) + \lambda^k(t) u_i(t) \\ \text{s.t. } & x_i(t+1) = A_i x_i(t) + B_i u_i(t) \\ \text{for } & t = s, s+1, \dots, T-1. \end{aligned}$$

In response, the ISO adjusts the prices according to: $\lambda^{k+1}(t) = \lambda^k(t)(1 - \alpha_k) + \alpha_k (\sum_i u_i^k(t))$, $t \geq s$. This process is iterated till it converges to $(u_i^*(s), u_i^*(s+1), \dots, u_i^*(T-1))$ and $(\lambda^*(s), \lambda^*(s+1), \dots, \lambda^*(T-1))$. At time s , the price is set at $\lambda^*(s)$ and agent i applies the input $u_i^*(s)$.

Algorithm 2

for bidding times $s = 0$ to $T-1$ **do**

$k = 0$

Initialize $\lambda^k(t), t \geq s$ to some arbitrary value.

repeat

Each agent i solves the problem

$$\min \sum_{t \geq s} x_i^\top(t) Q_i x_i(t) + u_i^\top(t) R_i u_i(t) + \lambda^k(t) u_i(t) \quad (9)$$

and submits the optimal values, denoted $u_i^k(t)$ for $t \geq s$ to the ISO.

ISO updates the prices,

$$\lambda^{k+1}(t) = \lambda^k(t)(1 - \alpha_k) + \alpha_k \left(\sum_i u_i^k(t) \right),$$

for times $t \geq s$.

Increment k by 1

until $u_i^k(t)$ converges to $u_i^*(t)$

implement $u^*(s)$

end for

The key to showing the existence of such a simple bidding scheme lies in utilizing the certainty equivalence property of LQG systems [26]. *Proof:* Let

$$x := (x_1, x_2, \dots, x_M), u := (u_1, u_2, \dots, u_M),$$

$$A := \text{diag}(A_1, A_2, \dots, A_M), B := \text{diag}(B_1, B_2, \dots, B_M),$$

$$Q = \text{diag}(Q_1, Q_2, \dots, Q_M), R = \text{diag}(R_1, R_2, \dots, R_M),$$

and consider the following deterministic linear, quadratic regulator (LQR) problem with no noise, but featuring the energy balance constraint,

$$\min \sum_{t=0}^T x^\top(t) Q x(t) + u^\top(t) R u(t) \quad (10)$$

$$\text{subject to } x(t+1) = Ax(t) + Bu(t), \quad (11)$$

$$\sum_{i=1}^M u_i(t) = 0 \text{ for } t = 0, \dots, T-1.$$

Since the state is affine in u , the cost is convex in u . Hence this centralized problem can be solved by the Bid-Price iteration between the agents and the ISO as shown for the deterministic problem. In particular, at time 0, the end result of the scheme is the optimal action $u(0)$. This is arrived at by the ISO announcing a sequence of prices for all future times and the agents bidding their consumptions/generation sequences at all future times. Now note that due to the energy balance at each time,

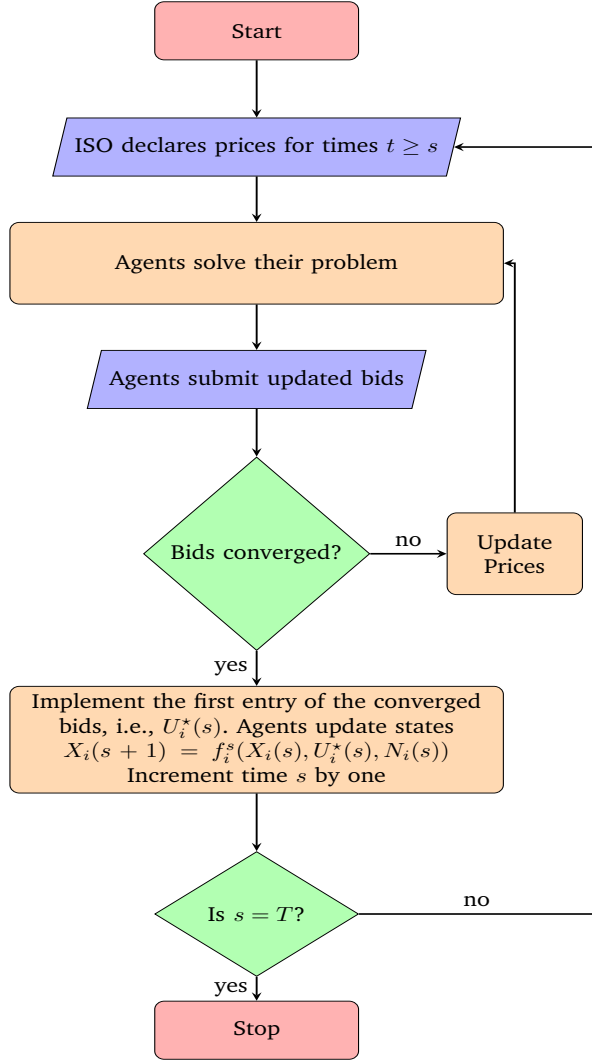


Fig. 2. Decision flow in Algorithm 1.

agent M is forced to choose $u_M(t) = -\sum_{i=1}^{M-1} u_i(t)$ for all t . Hence one can substitute this value for $u_M(t)$ and obtain a standard LQR problem where there is no separate energy balance constraint. For this reduced and standard deterministic linear quadratic regulator problem, the optimal solution is given by linear feedback $u(0) = \Gamma(0)x(0)$, where $\Gamma(\cdot)$ is the optimal feedback gain.

Now consider the corresponding reduced stochastic LQG problem where there is white Gaussian noise in the state equations (11). By Certainty Equivalence [26],

the same feedback law as in the deterministic reduced LQR problem is also optimal. In particular, in state $x(0)$ at time 0, $u(0) = \Gamma(0)x(0)$ continues to be optimal. Now, in our proposed bidding scheme for the LQG problem, each agent bids on the basis of a private deterministic system for itself. Hence it leads to the same Bid-Price iteration result at time 0. Hence it arrives at the same $u(0)$, which however is also optimal for the stochastic LQG problem.

Thus we see that the Bid-Price iteration scheme determines the optimal actions for the agents at time 0. Now our scheme for the LQG problem repeats such a Bid-Price scheme iteration at each time t . Each $x(t)$ can be regarded as an initial state for the system started at time t , and the same argument as above shows that the actions $u(t)$ that it results in for the agents at all times t are also optimal. ■

We note the following important aspects of the proposed algorithm. The critical feature that there is an iteration of bids at each time t is important. Also important is that at each stage it is the future sequence of prices that is iterated.

It should be noted that the alternative of announcing a “bid curve” of price vs. generation for a single time t does not work in the dynamic case. The reason is that the current optimal generation depends on future prices, so iteration of price at only one time is not sufficient to ensure optimal decisions when agents are dynamic systems.

XI. CONCLUDING REMARKS

We have posed the ISO Problem of maximizing the total utility/minimizing the total operating cost of the electricity grid, while not allowing the revelations of the dynamics, states or utilities of the agent. It is more complex than a decentralized stochastic control problem due to the non-revelation feature. We have shown that when the agents are LQG systems the problem admits a simple solution utilizing iterative bidding schemes, and attains the same performance as that of an optimal centralized control policy. Under the proposed policy, the sufficient statistics are vastly simplified, and each agent i needs to only keep track of its present state $X_i(t)$. This is in contrast to the general case of decentralized stochastic control, in which the agents need to keep track of the entire history in order to implement an optimal policy, which is in any case generally intractable to compute. Thus not only is our Algorithm decentralized, and easy to implement, but it also leads to a large amount of data reduction. We further note that our Algorithm is privacy preserving.

REFERENCES

- [1] F. Wu, K. Moslehi, and A. Bose, “Power system control centers: Past, present, and future,” *Proceedings of the IEEE*, vol. 93, no. 11, pp. 1890–1908, Nov 2005.

- [2] H.-P. Chao, S. Oren, A. Papalexopoulos, D. Sobajic, and R. Wilson, "Interface between engineering and market operations in restructured electricity systems," *Proceedings of the IEEE*, vol. 93, no. 11, pp. 1984–1997, Nov 2005.
- [3] F. Galiana, F. Bouffard, J. Arroyo, and J. Restrepo, "Scheduling and pricing of coupled energy and primary, secondary, and tertiary reserves," *Proceedings of the IEEE*, vol. 93, no. 11, pp. 1970–1983, Nov 2005.
- [4] K. J. Arrow, "An extension of the basic theorems of classical welfare economics," *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 507–532, 1951.
- [5] M. Morishima, *Walras' economics : a pure theory of capital and money*. Cambridge University Press, 1977.
- [6] H. S. Witsenhausen, "A counterexample in stochastic optimum control," *SIAM Journal on Control*, vol. 6, no. 1, pp. 131–147, 1968.
- [7] S. Yüksel and T. Basar, *Stochastic networked control systems: stabilization and optimization under information constraints*, ser. Systems & Control: Foundations & Applications. New York, NY: Springer, 2013.
- [8] E. Altman, *Constrained Markov Decision Processes*. Chapman and Hall/CRC, March 1999.
- [9] C. Striebel, "Sufficient statistics in the optimum control of stochastic systems," *Journal of Mathematical Analysis and Applications*, vol. 12, no. 3, pp. 576 – 592, 1965.
- [10] D. Blackwell and M. Girshick, *Theory of games and statistical decisions*. New York: John Wiley and Sons, 1954, republished by Dover in 1979. MR:0070134. Zbl:0056.36303.
- [11] H. Gottinger, W.-D. Eberwein, and R. Bogdan, "Reviews," *Theory and Decision*, vol. 5, no. 3, pp. 349–354, 1974. [Online]. Available: <http://dx.doi.org/10.1007/BF00164728>
- [12] A. M. Ashutosh Nayyar and D. Teneketzis, "Chapter 4: The Common-Information Approach to Decentralized Stochastic Control," in *Information and Control in Networks*. Springer-Verlag, 2014, pp. 123–156.
- [13] J. Wu, "Sufficient statistics for team decision problems," Ph.D. dissertation, Stanford University, November 2013.
- [14] K. J. Arrow, "General Economic Equilibrium: Purpose, Analytic Techniques, Collective Choice," *American Economic Review*, vol. 64, no. 3, pp. 253–72, June 1974.
- [15] Fred C. Schweppe, Michael C. Caramanis, Richard D. Tabors, Roger E. Bohn, *Spot Pricing of Electricity*. Springer US, 1988.
- [16] M.C. Caramanis, R.E. Bohn, and F.C. Schweppe, "Optimal spot pricing: Practice and theory," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-101, no. 9, pp. 3234–3245, Sept 1982.
- [17] M. D. Ilić, J.-Y. Joo, L. Xie, M. Prica, and R. N., "A decision-making framework and simulator for sustainable electric energy systems," *IEEE Transactions on Sustainable Energy*, vol. 2, no. 1, pp. 37–49, Jan. 2011.
- [18] Q. Zhu, P. Sauer, and T. Basar, "Value of demand response in the smart grid," in *IEEE Power and Energy Conference at Illinois (PECI)*, Feb. 2013, pp. 76–82.
- [19] Y. Wang, W. Saad, Z. Han, H. V. Poor, and T. Basar, "A game-theoretic approach to energy trading in the smart grid," *IEEE Transactions on Smart Grid*, vol. 5, no. 3, pp. 1439–1450, May 2014.
- [20] V. S. Borkar, *Stochastic approximation : A dynamical systems viewpoint*. Cambridge: Cambridge University Press New Delhi, 2008.
- [21] D. P. Bertsekas, A. E. Ozdaglar, and A. Nedic, *Convex analysis and optimization*, ser. Athena scientific optimization and computation series. Belmont (Mass.): Athena Scientific, 2003.
- [22] R. Tyrrell Rockafellar, *Convex Analysis*. Princeton University Press, 1996.
- [23] S. Boyd. [Online]. Available: <https://web.stanford.edu/class/ee392o/subgrad.method.pdf>
- [24] H. J. Kushner and G. Yin, *Stochastic Approximation Algorithms and Applications*. New York: Springer Verlag, 1997.
- [25] H. Robbins and S. Monro, "A stochastic approximation method," *The Annals of Mathematical Statistics*, vol. 22, no. 3, pp. 400–407, Sept. 1951.
- [26] P. R. Kumar and P. Varaiya, *Stochastic systems: Estimation, identification and adaptive control*. Prentice Hall Inc., Englewood Cliffs, 1986.